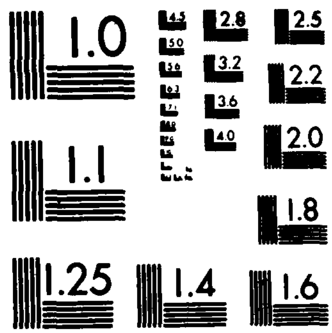
AD-A191 915 EMPIRICAL MODELS REPRESENTING THE ERROR IN THE PREDICTED MUF AND FIELD STRENGTH FROM HFBC84(U) NAVAL OCEAN SYSTEMS CENTER SAN DIEGO CA D B SAILORS NOV 87 1/1

UNCLASSIFIED F/G 20/14 NL


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

AD-A191 915

DTIC FILE COPY

UNCLASSIFIED
SECURITY CLASSIFICATION OF

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Ocean Systems Center 6b. OFFICE SYMBOL (if applicable) NOSC	7a. NAME OF MONITORING ORGANIZATION Naval Ocean Systems Center
6c. ADDRESS (City, State and ZIP Code) San Diego, CA 92152-5000	7b. ADDRESS (City, State and ZIP Code) San Diego, CA 92152-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Space and Naval Warfare Systems Command 8b. OFFICE SYMBOL (if applicable) SPAWAR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code) REWSON Systems Program Washington, DC 20363	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO.: RDDA PROJECT NO.: MP57 TASK NO.: NSA AGENCY ACCESSION NO.: DN288 591

11. TITLE (include Security Classification)

Empirical Models Representing the Error in the Predicted MUF and Field Strength from HFBC 84

12. PERSONAL AUTHOR(S)

D.B. Sailors

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Presentation Speech	FROM May 1987 TO May 1987	November 1987	

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	diurnal trends; sunspot numbers; geographic region

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Models of the probability distribution representing the residuals between the observed MOF and field strength and the corresponding predicted values from HFBC84, an ionospheric prediction program developed for the Broadcast WARC, are presented. A data base of 13,054 hours of oblique sounder MOFs measured on 70 paths was used to obtain the residuals in the predicted MUFs. A modified version of CCIR Data Base C was used to obtain the residuals in predicted field strength; only the 81 paths for short path propagation were retained. The residuals for these models were fit to the Johnson system of frequency curves. This system of curves consists of three distributions: (1) data that is unbound, called the S_U distribution; (2) data that is bound on one end called the S_L distribution; and (3) data that is bound on both ends, called S_B distribution. Overall for the MUF residuals, it was determined that the data could be represented by a S distribution. Overall for the field strength residuals, a S_U distribution fit the data for frequencies below or equal the predicted MUF, and a S_B distribution fit the data for frequencies greater than MUF.

DTIC SELECTE MAR 2 8 1988 S D

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL D.B. Sailors	22b. TELEPHONE (include Area Code) (619)225-7706 22c. OFFICE SYMBOL Code 542

DD FORM 1473, 84 JAN

83 APR EDITION MAY BE USED UNTIL EXHAUSTED
ALL OTHER EDITIONS ARE OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

EMPIRICAL MODELS REPRESENTING THE ERROR IN THE PREDICTED MUF AND FIELD STRENGTH FROM HFBC84

D. B. SAILORS
OCEAN AND ATMOSPHERIC SCIENCES DIVISION
NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CA 92152-5000

ABSTRACT

Models of the probability distribution representing the residuals between the observed MOF and field strength and the corresponding predicted values from HFBC84, an ionospheric prediction program developed for the Broadcast WARC, are presented. A data base of 13,054 hours of oblique sounder MOFs measured on 70 paths was used to obtain the residuals in the predicted MUFs. A modified version of CCIR Data Base C was used to obtain the residuals in predicted field strength; only the 81 paths for short path propagation were retained. The residuals for these models were fit to the Johnson system of frequency curves. This system of curves consists of three distributions: (1) data that is unbound, called the S_U distribution; (2) data that is bound on one end, called the S_L distribution; and (3) data that is bound on both ends, called S_B distribution. Overall for the MUF residuals, it was determined that the data could be represented by a S_L distribution. Overall for the field strength residuals, a S_U distribution fit the data for frequencies below or equal the predicted MUF, and a S_B distribution fit the data for frequencies greater than the MUF.

INTRODUCTION

The effective operation of long distance high frequency (HF) communication systems has increased in proportion to the ability to predict variations in the ionosphere. These variations are affected in a complex manner by solar activity, seasonal and diurnal changes, as well as latitude and longitude. Such a predictive capability has permitted communicators to optimize frequencies, antennas and other circuit parameters. The need for HF model uncertainty assessment has become important as more and more uses are found for the prediction models. This paper differs from the more usual report on the accuracy of prediction models. Here, a model of the probability distributions representing the residuals (the errors) between the observed parameters and the corresponding predicted values is presented. The residuals for the prediction model are fit to a Johnson system of frequency curves (Johnson, 1949) using an algorithm due to Hill et al. (Hill et al., 1976; Hill and Wheeler, 1981; Dodgson and Hill, 1983) which uses the method of moments to obtain the required parameters. This distribution represents all univariate distribution systems; its simplicity of calculation once its parameters have been determined makes it adaptable to minicomputer type of applications; and the transformation of the Johnson variables to the normal system allows use of normal probability algorithms in its application.

Here, the probability distribution representing the error in the predicted MUF and field strength from HFBC84, an ionospheric prediction program for the HF Broadcasting WARC, is presented. A data base of 13,054 hours of oblique sounder MOFs measured on 70 paths was used to obtain the residuals in the predicted MUFs. A modified version of CCIR Data Base C was used to obtain the residuals in predicted field strength; only the 81 paths for short path propagation were retained. Detailed documentation of the determination of the residuals for the predicted MUF asnd field strength for HFBC84 is given by Sprague and Sailors (1987), Roy and Sailors (1987), and Sprague (1987).

The first section of the body of the paper discusses the determination of the statistical model. In addition to the method of moments, this section presents alternate methods for determining the parameters when the moments are large. The second section presents tables of parameters representing the probability distribution of HFBC84 predictions. However, because of the limited scope of the paper, not all the variations studied can be presented. The third section discusses application of the statistical model. The final section discusses improvements that could be made in the technique used to get the Johnson curve parameters.

DETERMINATION OF A STATISTICAL MODEL

The first step in determining a probability distribution to represent the residuals between observed data and the corresponding predicted data is to hypothesize a distribution function and determine its parameters. In the case of the application at hand, the determination of the distribution function was motivated by (1) a desire for ease of calculation so that the results could be used in minicomputer type applications, and (2) a need to automate the data analysis. The distribution function chosen was one known as the Johnson curves (Johnson, 1949).

Johnson Curves

The Johnson curves are an empirical family of curves satisfying the following chosen conditions: (1) they should be easy to evaluate once their parameters are determined; (2) they are a monotonic function of y where $y = (x-\xi)/\lambda$, λ is a scale factor, ξ a location factor of the distribution, and x is the variable being represented by the distribution; (3) the range of values of f(y) corresponding to the actual range of values of y should be from $-\infty$ to $+\infty$; and (4) the resulting system of distributions of y (and so of x) should include distributions of most, is not all, of the kinds encountered in collected data. Johnson proposed basing empirical distributions on the transformation of a standard normal variate. The Johnson system of frequency curves consist of:

the lognormal system (or S_L): $z = \gamma + \delta \,\ln [(x-\xi)/\lambda]$, $\xi<x$, (1)

the unbounded system (or S_U): $z = \gamma + \delta \sinh^{-1} [(x-\xi)/\lambda]$, (2)

the bounded system (or S_B): $z = \gamma + \delta \,\ln [(x-\xi)/(\xi+\lambda-x)]$, $\xi<x<\xi+\lambda$ (3)

where z is the standardized normal variate in each case. The parameters γ and δ determine the shape of the distribution of x.

To decide which of the three Johnson families should be used for a given set of data, the usual procedure is to obtain the data estimates of the skewness $\sqrt{\beta_1}$ and the kurtosis β_2. These then are plotted on a figure such as figure 1 which shows the region in the (β_1, β_2) plane for the three Johnson families. Also shown are other common sampling distributions. If the (β_1, β_2) point is close to the S_L curve, the S_L curve is chosen. If it is in the region above the S_L curve, the S_B family is chosen; and if it is below the curve, the S_U curve is used. The S_L curve can be extended by use of the parametric equations:

$$\beta_1 = (\omega-1)\,(\omega+2)^2 \tag{4}$$

$$\beta_2 = \omega^4 + 2\,\omega^3 + 3\,\omega^2 - 3 \tag{5}$$

where ω denotes $\exp(\delta^{-2})$. The impossible region in the figure is bounded by the line $\beta_2 - \beta_1 - 1 = 0$. These three systems of curves; S_L, S_U, and S_B; together cover the entire "possible" region of the β_1, β_2 plane (the plane describing the entire possible 3rd and 4th moment variation).

Fitting Johnson Curves by Moments

To determine the parameters for the Johnson curves, an algorithm, called JNSN, known by the Royal Statistical Society as algorithm AS 99, was used (Hill et al., 1976; Hill and Wheeler, 1981; Dodgson and Hill, 1983; Griffiths and Hill, 1985). This algorithm uses the sample moments (i.e., the mean, standard deviation, skewness ($\sqrt{\beta_1}$), and the kurtosis (β_2) to determine the type of Johnson curve and its parameters. For the sake of completeness, the algorithm includes the normal curve itself and the special case of the S_B curve on the $\beta_2 = \beta_1 + 1$ boundary which is called S_T (T standing for two-ordinate). Equations (4) and (5) are solved in JNSN to determine the type of Johnson distributions. If the required β_2 from these equations is less than the input β_2, S_B (or S_T) is appropriate; if greater, S_U is appropriate. The parameters γ, δ, ξ, and λ are found such that estimated values of $\sqrt{\beta_1}$ and β_2 using the parameters are within ± 0.01 of the input values.

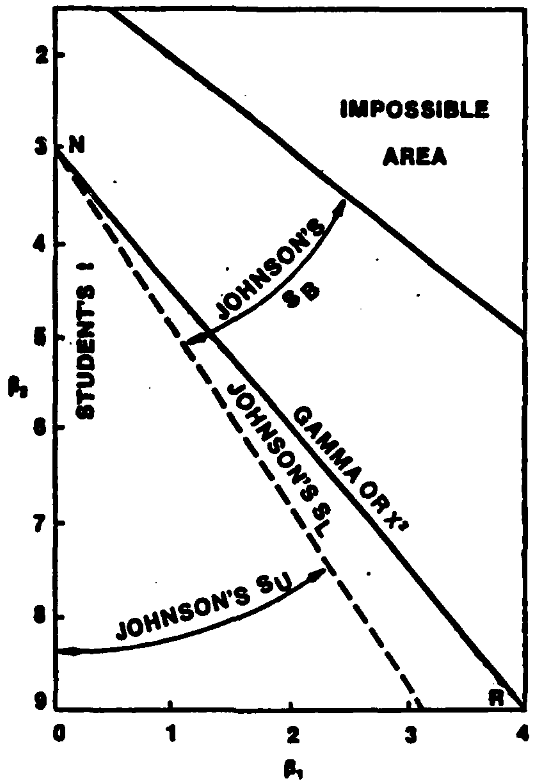


Figure 1. Region in (β_1, β_2) plane for the Johnson System of Curves.

Alternate Methods for Determining the Parameters

When the moments are large, the methods of moments is not always efficient. Hence, it is desirable to have alternate methods to use to obtain a second set of parameters. A test of fit can be used to choose between the sets of parameters so chosen. In addition to the methods of moments, two additional methods of estimation are maximum likelihood and the use of quantiles or percentiles. Depending on the Johnson curve being evaluated, the different methods may or may not be possible.

For the S_L curve, the maximum likelihood method can be used to obtain the parameters γ and δ if β is known (Hahn and Shapiro, 1967). If ξ is unknown, one is tempted to set ξ to the minimum value in the data set. However, as ξ tends to the minimum, the maximum likelihood estimates tend to infinity. There can be considerable variation of ξ with little variation of the percentiles. This is particularly true for large negative values of ξ. In the algorithm written for this method, ξ was set to min $(x_1, x_2, \ldots, x_N)$ - 0.1 min $(x_1, x_2, \ldots, x_N)$. If the value of ξ is unknown, the unknown parameter γ, δ, and ξ can be obtained by setting three percentiles determined from the data to the three corresponding percentiles for the normal standard variate z (Hahn and Shapiro, 1967). Three equations of the form

$$z_\alpha = \hat{\gamma} + \hat{\delta} \ln (x_\alpha - \hat{\xi}) \tag{6}$$

are solved for the parameters $\hat{\gamma}$, $\hat{\delta}$, and $\hat{\xi}$; x_α is the α 100th percentile for a standard normal variate and x_α is the corresponding percentile from the data. Hahn and Shapiro provide solutions to these equations for any two symmetric percentiles (e.g., the α 100th and the $(1-\alpha)$ 100th percentile) along with the median value. If $(x_{1-\alpha} - x_{0.5}) < (x_{0.5} - x_\alpha)$, then a negative value of δ is obtained and the percentile method fails; this is very unlikely to be the case if the distribution has substantial positive skewness.

For the S_U case, the only alternate method is the method of quantiles. In this case all four parameters, ξ, λ, γ, and δ, have to be estimated. This requires that estimates be obtained of four values, x_A, x_B, x_C, x_D say, such that certain fixed proportions P_A, P_B, P_C, P_D, respectively, of the distribution fall below these values. Then the parameters may be found from the four equation for z_k, the normal standard deviate for the S_U case. This set of equations is non-linear in nature and currently there is no algorithm that has been used to solve them.

For the S_B curves, both the maximum likelihood and method of percentile methods can be applied (Johnson, 1949). If the parameters ξ and λ are known, then the maximum likelihood method can be used to obtain the parameters γ and δ. If only the lower end-point, ξ, is known, the method of percentiles can be used to estimate λ, γ, and δ. The method of percentiles uses the medium x_o and the lower and upper 100th percentile points x_1 and x_2.

Verification of Model Fit

The final step in the determination of a distribution representing the error data was the verification that the fit was adequate. In addition the test of fit was used to determine if the parameters produced by alternate methods were preferable. The procedures followed for the test of fit was that for the chi-square test of fit (Hahn and Shapiro, 1967; Williams, 1950). The first step is the estimation of the unknown parameters of the assumed distribution (Johnson curves). After sorting the data, the data is divided into k classes or cells, and the probability of a random value from the assumed model falling with each class is determined. The number of classes k itself is selected by the following formula, depending on the sample size N and the level of significance C

$$k = 4\,[2(N-1)^2/C^2]^{1/5}. \tag{7}$$

For the 5% level of significance C = 1.645. The class limits are chosen such that each class contains the same number of items under the null hypothesis (i.e., a Johnson distribution). For this distribution, the class boundaries were found by inputting 1/k, 2/k, ..., (k-1)/k into an algorithm for finding the normal deviates corresponding to the lower tail (Beasley and Springer, 1977). Then an algorithm was used to find the Johnson deviates corresponding to the normal deviates (Hill and Wheeler, 1981; Dodgson and Hill, 1983; Hill, 1976). These Johnson deviates so determined are the cell boundaries. The lower boundary of the first cell and the upper boundary of the last cell are the smallest and largest values that the random variable may take on. The cell boundaries were set up in such a way that the probability of a random value falling within a given class is estimated to be 1/k for each class. The expected number of observations for each cell under the assumed distribution is N/k. The chi-square test statistic is given by

$$\chi^2 = \frac{k}{N} \sum_{i=1}^{k} f_i^{\,2} - N \tag{8}$$

where f_i is the observed number of frequencies in the ith class. As the sample size approaches infinity, the distribution of this statistic approaches the chi-square distribution with k-1-s degrees of freedom where s is the number of parameters estimated from the sample.

For a one-sided (upper tail) test for a significance level of 5%, $\chi^2_{k-1-s}(\alpha)$ is determined using a table of critical values for the chi-square distribution for degrees of freedom less than equal ten and the Cornish-Fisher approximation for larger degrees of freedom (Goldberg and Levine, 1945; Zar, 1978). Ratios of $\chi^2/\chi^2_{k-1-s}(\alpha)$ greater than 1 signify that the observed data contradicts the assumed model. This ratio is also used to determine which of the methods of determining the Johnson parameters is preferable; the method producing the lowest ratio is used. The chi-square test holds for sample sizes as low as 200 (Williams, 1950).

The chi-square test depends on the data being ungrouped. When data is grouped, the chi-square test will fail. That is because equations (7) and (8) assume that there are N ungrouped data points. What happens is that χ^2 calculated from equation (8) is large and the hypothesis is rejected when in fact it might be accepted. This is because χ^2 is computed on the assumption that every cell has an expected number not too small. Usually 5 is given as the required minimum. When the data is grouped, then some cells have large numbers in them and adjoining cells have a number less than the minimum or even zero. One solution is to include cells with numbers less than the minimum with adjoining cells, and increase the number of cells. However, the number of cells must not exceed N/5. In the application here the number of cells was doubled when grouping occurred.

JOHNSON CURVE PARAMETERS FOR HFBC84

The Johnson curve parameters were determined for the predicted MUF and field strength for HFBC84. For the predicted MUF, the parameters were found as a function of latitude of the control points, of data recorder type, of geographics region, of each particular path in the data base where sufficient data was available, of circuit path length, of mid-path local time, of monthly mean sunspot number, of month, of path orientation, of season, and of smoothed sunspot number. A limited set of these results are presented in tables 1 to 6. For the predicted field strength, the parameters were found as a function of frequency relative to predicted MUF, of each particular path in the data base where sufficient data was available, of mid-path latitude, of circuit length, of mid-path local time, of relative path orientation, of season, of smoothed sunspot number, and of mid-path local time. In the case of the field strength, the parameters were determined for frequencies > MUF and frequencies ≤ MUF. A limited set of these results is presented in tables 7 to 14. In the case of the field strength data, the alternate methods produced the preferable set of parameters quite often, particularly the maximum likelihood method. Whereas, for the MUF residuals, only rarely did an alternate method to the method of moments produce the preferable set of parameters.

In tables 1 to 14, the conditions for which the analysis was made (e.g., circuit length 2000-3000 km), the sample size, the four input moments, γ, δ, λ, ξ, the χ^2/χ^2_{k-1-s} (0.05) ratio, the class size k, the Johnson type curve t, and the method of analysis m are given. Type = 1 implies an S_L curve, type = 2 implies an S_U curve; type = 3 implies an S_B curve; type = 4 implies a normal curve; and type = 5 implies an S_T curve. Method m = 1 implies method of moments; method m = 2 implies percentile points; and method m = 3 implies maximum likelihood. In the table the error codes imply: (1) 1 = residual statistics were not calculated; (2) 2 = residual data is not valid; (3) 3 = chi-squared test of fit unable to run, usually not enough data points; and (4) 4 = chi-squared critical value not determined usually not enough degrees of freedom. Only the error codes 3 and 4 having to do with the ch-square test of fit actually occurred. This usually was due to grouping in the residual data, particularly the field strength data. Because the data in Date Base C is given only as integer values, this causes grouping precisely at values of field strength where the model is most accurate.

APPLICATION OF THE MODEL

After having determined the Johnson distribution parameters for HFBC84 by the methods described earlier, it is desirable to be able to apply the results. It is envisioned that it might be applied in one of two ways. The first is that for a given probability, it might be desired to know the error in the model. The second is given a certain error in the model, what is the corresponding probability? There are several useful algorithms to aid in this (Beasley and Springer, 1977; Dodgson and Hill, 1983; Griffiths and Hill, 1985; Hill, 1973; Hill, 1976; Hill and Wheeler, 1981).

In the first application, the given probability is converted to the corresponding normal standard deviate using the algorithm function PPND (Beasley and Springer, 1977; Griffith and Hill, 1985). Then the corresponding Johnson deviates are found using the algorithm AJV (Dodgson and Hill, 1983; Griffiths and Hill, 1985; Hill, 1976; Hill and Wheeler, 1981). The parameters necessary as input can be found for the case and model under consideration from the tables in the previous section. The Johnson deviates are the error for the model being applied.

In the second application, the given error is converted to normal standard deviates using the second algorithm due to Hill (1976; Dodgson and Hill, 1983; Griffiths and Hill, 1985; Hill and Wheeler, 1981) called SNV. Then the corresponding probability level can be found using a normal integral algorithm (Hill, 1973) called ALNORM. This particular algorithm has the capability to calculate either the upper or lower tail area of the standardized normal curve corresponding to any given argument.

A sample of the application of the model is given in figures 2-3. In these figures the predicted residual is given for seven different standard normal deviates (snv) and their corresponding probability levels. The residuals range from values that might occur 0.1% of the time to 99.9% of the time. Figure 2 is for the MUF model, and shows the residual variation as a function of path range. Figure 3 is for the field strength model, and shows the residual variation as a function of ratio of the frequency to the predicted MUF.

TABLE 1. Johnson Curve Parameters for HFBC84 MUF : Circuit Length (x1000 km)

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
0-1	396	1.90	1.71	.493	3.575	3.002	2.215	22.940	-3.02	.996	45	3	1
1-2	737	2.59	2.08	.278	6.630	-.295	1.578	2.593	2.00	2.039	62	2	1
2-3	1014	1.51	3.89	2.798	8.986	2.154	1.838	3.300	7.10	2.249	88	2	1
3-4	1254	2.26	3.32	.187	7.523	.193	1.467	3.752	2.89	2.207	88	2	1
4-5	1738	-3.88	4.32	.004	3.074	1.286	7.734	32.680	1.62	1.594	130	2	1
5-6	1581	.78	4.67	.024	3.801	-.297	2.572	11.040	-.60	1.707	128	2	1
6-7	1291	-.40	4.44	.140	2.805	-1.102	1.783	36.550	-23.80	1.021	124	3	1
7-8	1054	3.20	3.99	.147	2.683	-.863	1.490	27.710	-14.21	1.729	102	3	1
8-9	1169	4.45	2.05	.395	4.344	1.275	2.531	4.195	6.83	1.623	74	2	1
9-10	1630	2.28	4.13	2.059	6.285	-4.973	2.337	-1.000	11.49	4.291	105	1	1
10-11	1190	1.61	4.18	1.148	4.882	-5.640	2.325	120.500	-108.40	3.670	87	3	1

TABLE 2. Johnson Curve Parameters for HFBC84 MUF : Season

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
WINTER	4252	1.09	4.52	.370	3.774	4.137	4.009	11.000	14.99	4.044	151	2	1
SPRING	2232	.04	4.88	.192	3.860	1.129	2.850	12.060	5.26	3.742	139	2	1
SUMMER	4213	2.24	3.96	1.386	5.302	-5.634	2.192	119.700	-108.30	3.378	132	3	1
FALL	2357	.45	4.66	.383	3.583	-6.040	3.347	125.900	-107.20	2.822	138	3	1

TABLE 3. Johnson Curve Parameters for HFBC84 MUF : Latitude of Control Points

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
Transequat.	2631	.30	5.14	.098	3.649	.902	3.032	14.090	4.79	2.280	151	2	1
Low Latitude	1114	-1.67	4.62	.157	4.101	.686	2.433	9.873	1.39	1.372	110	2	1
Mid Latitude	4350	-.06	4.66	.280	3.085	-1.710	1.897	44.030	-30.89	4.880	156	3	1
High Latitude	4045	3.43	2.85	.367	5.486	.512	1.832	4.276	4.83	2.205	108	2	1
Transauroral	914	2.92	2.31	.070	4.524	-.290	2.040	4.113	2.25	1.295	68	2	1

TABLE 4. Johnson Curve Parameters for HFBC84 MUF : smoothed SSN

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
0 - 30	2649	.70	4.49	.698	4.399	3.904	3.224	7.414	12.59	2.734	138	2	1
31 - 60	1512	1.99	3.61	.146	5.063	.333	1.865	5.692	3.17	2.803	99	2	1
61 - 90	1908	1.38	4.14	.007	4.231	-.106	2.155	7.972	.95	2.641	125	2	1
91 - 120	6469	1.01	4.89	.679	3.354	-1.425	1.225	33.680	-23.75	5.236	165	3	1
121 - 150	516	2.26	2.52	.008	4.702	.085	1.924	4.205	2.47	1.081	54	2	1

TABLE 5. Johnson Curve Parameters for HFBC84 MUF : mid path local time

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
1	566	1.30	3.61	.332	3.244	-2.205	2.097	39.820	-27.81	.836	65	3	1
2	529	1.05	3.46	.145	3.079	-2.712	3.014	51.240	-35.14	1.482	60	3	1
3	540	1.01	3.23	.042	3.271	1.380	4.480	13.450	5.32	1.219	62	2	1
4	532	1.48	2.75	.060	3.499	.856	3.320	8.436	3.78	.997	61	2	1
5	539	2.09	2.66	.052	6.374	-.120	1.568	3.347	1.78	1.796	53	2	1
6	542	2.18	3.00	.808	4.286	-5.557	2.578	78.820	-68.02	1.589	59	3	1
7	537	2.04	3.54	.895	4.363	-4.556	2.287	72.530	-61.23	1.503	58	3	1
8	544	1.71	4.11	.410	3.120	-1.450	1.501	31.940	-20.87	1.065	60	3	1
9	544	1.58	4.79	.383	3.830	3.583	3.783	11.680	14.83	1.535	56	2	1
10	543	1.83	4.68	.534	3.470	-2.101	1.756	46.140	-32.94	1.946	50	3	1
11	548	1.88	4.47	.663	3.592	-1.993	1.577	40.800	-29.22	1.245	59	3	1
12	548	1.54	4.81	.473	3.504	-2.686	2.108	58.940	-43.92	1.205	59	3	1
13	550	1.14	5.16	.344	3.353	-2.898	2.421	69.910	-52.01	1.100	56	3	1
14	545	.64	5.48	.314	3.416	-4.509	3.120	109.400	-87.40	1.037	58	3	1
15	543	.30	5.62	.400	2.797	-.949	1.083	31.880	-21.32	1.170	57	3	1
16	544	.26	5.63	.462	2.756	-.884	.946	29.010	-19.53	1.282	56	3	1
17	548	.37	5.54	.398	2.619	-.787	.902	27.130	-17.76	1.547	54	3	1
18	547	.48	5.42	.311	2.441	-.663	.851	24.920	-15.68	.998	59	3	1
19	551	.38	5.30	.318	2.786	-.977	1.221	32.670	-21.45	.982	57	3	1
20	551	.67	5.04	.377	4.018	1.982	3.060	11.890	9.37	1.255	53	2	1
21	548	.98	4.63	.315	5.631	.429	1.772	6.713	2.91	1.165	55	2	1
22	550	1.04	4.25	.502	5.698	.600	1.826	6.243	3.46	1.354	57	2	1
23	545	1.02	3.97	.836	4.740	3.108	2.861	6.340	9.87	1.107	61	2	1
24	520	1.03	3.85	.367	4.023	1.940	3.015	8.969	7.56	.798	59	2	1

TABLE 6. Johnson Curve Parameters for HFBC84 MUF : Overall Analysis

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
all data	13054	1.16	4.52	.565	4.013	-11.910	4.132	-1.000	19.56	6.205	173	1	1

TABLE 7. Johnson Curve Parameters for HFBC84 Field Strength : Overall Analysis

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
all data	12277	3.97	16.20	.244	3.626	-2.849	3.873	46.980	-35.06	9.517	94	2	1
freq<=MUF	8783	1.93	11.91	.032	3.611	-.460	2.912	32.230	-3.49	5.732	71	2	1
freq >MUF	3494	9.10	23.00	.002	2.133	.067	.989	109.600	-44.17	3.438	94	3	1

Start here for all pages after the first

On first page type title of paper here

Author
Affiliation
City, State

TABLE 8. Johnson Curve Parameters for HFBC84 Field Strength : freq/MUF ratio

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
<=.2	41	-1.10	10.66	.084	3.911	-.545	2.536	24.360	-6.80	* error = 3		2	1
.3	355	1.83	12.79	.245	5.343	-.407	1.822	19.340	-3.24	1.516	35	2	1
.4	1061	2.40	12.52	.016	2.933	.964	1.502	92.400	-30.80	2.734	52	3	3
.5	1446	2.27	12.02	.063	2.898	.415	1.417	78.100	-31.90	2.362	53	3	3
.6	1555	3.28	11.48	.105	3.056	3.108	3.546	197.500	-55.42	1.868	51	3	1
.7	1348	3.84	10.77	.368	5.902	-.440	1.726	14.990	-.73	1.582	47	2	1
.8	1166	2.43	11.23	.030	3.305	-.941	4.072	43.190	-7.94	1.552	50	2	1
.9	1031	-.18	11.53	.000	3.177	.145	4.907	55.380	1.49	1.140	49	2	1
1.0	780	-3.06	12.88	.003	3.115	.645	6.104	77.160	5.23	1.521	50	2	1
1.1	668	-5.96	15.05	.015	3.018	-144.200	24.420	-1.000	361.40	1.388	55	1	1
1.2	558	-8.10	13.88	.040	3.177	-2.372	5.808	73.130	-39.27	1.098	47	2	1
1.3	347	-8.47	14.73	.051	3.640	-.579	2.904	39.440	-16.87	1.295	35	2	1
1.4	277	-6.81	13.17	.000	3.333	.081	3.675	46.600	-5.74	.826	29	2	1
1.5	214	1.52	12.91	.614	4.194	4.832	3.549	20.840	41.05	.881	22	2	1
1.6	175	14.53	11.02	.000	3.758	.000	2.586	26.390	14.53	1.288	19	2	1
1.7	142	28.66	11.00	.072	2.613	.666	1.595	78.840	-3.30	.829	12	3	1
1.8	106	40.96	10.38	.117	3.530	1.387	3.495	32.140	54.60	.738	10	2	1
1.9	120	42.37	9.67	.119	2.525	-.644	1.292	58.700	6.66	1.309	8	3	1
2.0	86	39.63	8.49	.222	3.691	1.936	3.422	23.770	54.43	1.866	6	2	1
2.3	356	36.31	9.20	.042	2.465	.475	1.037	47.300	16.20	.846	34	3	3
2.8	187	31.62	7.81	.618	4.385	-2.642	2.973	15.190	15.37	.808	20	2	1
3.3	109	28.50	5.83	.005	3.040	-3.095	11.070	61.800	10.93	.294	12	2	1
3.8	74	24.96	3.88	1.016	4.481	3.831	2.021	65.000	15.82	5.224	6	3	1
4.3	37	21.57	6.41	.000	2.621	.038	1.960	53.310	-4.84	* error = 3		3	1
4.8	23	20.43	4.86	.002	1.797	-.050	.635	17.480	11.45	* error = 3		3	1
>5.0	15	18.40	5.58	.032	3.490	-.574	3.222	16.850	15.23	* error = 3		2	1

TABLE 9. Johnson Curve Parameters for HFBC84 Field Strength : smoothed SSN with freq > MUF

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
0 - 30	717	5.52	20.74	.029	2.066	.031	1.257	117.700	-50.60	4.036	59	3	3
31 - 60	679	5.76	23.61	.000	2.058	-.015	.905	105.700	-47.46	1.637	74	3	1
61 - 90	525	7.35	24.14	.001	2.153	-.060	1.015	117.200	-52.66	1.482	59	3	1
91 - 120	694	15.24	23.11	.003	1.988	-.066	.827	97.490	-35.01	1.196	74	3	1
121 - 150	495	12.75	20.26	.018	2.428	.276	1.399	126.900	-45.09	1.376	55	3	1
151 - 180	384	8.24	24.49	.049	1.885	.060	.893	106.700	-41.80	1.994	37	3	3

TABLE 10. Johnson Curve Parameters for HFBC84 Field Strength : smoothed SSN with freq <= MUF

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
0 - 30	1279	-.47	11.09	.000	3.282	-.099	3.956	42.470	-1.57	1.252	50	2	1
31 - 60	2096	1.33	11.69	.002	3.528	-.135	3.011	33.230	-.24	1.491	56	2	1
61 - 90	1948	2.37	11.91	.002	2.915	-.667	4.899	236.800	-124.00	1.618	54	3	1
91 - 120	1295	3.06	10.78	.046	3.630	-.554	2.913	29.000	-2.83	1.776	45	2	1
121 - 150	1343	3.45	12.98	.163	4.381	-.544	2.201	24.840	-3.43	3.112	55	2	1
151 - 180	822	1.91	12.84	.225	3.266	4.188	3.320	246.800	-53.62	2.040	47	3	1

Leave Blank

TABLE 11. Johnson Curve Parameters for HFBC84 Field Strength : Circuit Length (x1000 km) with freq > MUF

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
0-1	2038	15.64	22.96	.024	1.799	-.148	.614	81.380	-28.38	4.329	81	3	1
1-2	282	13.11	19.24	.001	1.880	-.031	.727	74.780	-24.86	3.346	22	3	1
3-4	82	-26.70	14.73	.041	2.452	.408	1.375	91.720	-66.50	* error = 4		3	1
5-6	172	-11.52	17.56	1.335	4.509	2.166	1.313	159.400	-41.06	1.196	12	3	1
6-7	175	-17.52	16.63	1.448	6.520	-1.598	2.018	21.630	-38.98	1.922	16	2	1
7-8	6	.52	13.71	.733	2.683	.768	.602	54.140	-15.50	* error = 3		3	1
9-10	474	9.66	12.75	.003	4.120	-.072	2.228	25.590	8.75	1.726	41	2	1
10-11	24	-.46	13.83	2.691	6.164	2.048	.972	113.300	-16.92	* error = 3		3	1
16-17	240	-4.78	10.13	.016	2.714	-.486	2.224	95.360	-57.41	.741	25	3	1

TABLE 12. Johnson Curve Parameters for HFBC84 Field Strength : Circuit Length (x1000 km) with freq <= MUF

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
0-1	2132	-4.87	8.76	.001	2.534	.079	1.694	64.090	-36.22	2.885	40	3	1
1-2	904	-.46	7.92	.002	2.960	-406.800	65.130	-1.000	515.60	1.313	35	1	1
2-3	143	.06	9.19	.002	3.753	-.082	2.596	22.110	-.69	2.689	12	2	1
3-4	625	8.32	15.26	.046	3.663	.523	2.844	40.030	16.19	1.961	48	2	1
4-5	114	3.37	10.20	.002	2.914	.519	4.501	186.500	-84.57	.522	9	3	1
5-6	1483	2.09	9.51	.321	4.049	1.497	2.825	21.870	15.02	3.981	43	2	1
6-7	937	.17	13.44	.044	2.948	-2.109	3.683	217.400	-138.30	1.269	55	3	1
7-8	65	15.00	8.41	.578	2.716	-.820	.771	38.020	-11.30	* error = 4		3	1
8-9	37	16.35	8.07	.191	2.111	-.447	.701	31.890	-3.22	* error = 3		3	1
9-10	1539	10.75	11.14	.003	4.234	.069	2.150	21.390	11.52	1.641	49	2	1
10-11	41	12.93	16.16	.016	1.743	.117	.578	55.400	-12.93	* error = 3		3	1
12-13	12	19.83	7.31	.616	2.620	-.754	.662	30.040	-.95	* error = 3		3	1
16-17	751	.04	8.95	.005	4.115	.099	2.234	18.010	.93	1.780	36	2	1

TABLE 13. Johnson Curve Parameters for HFBC84 Field Strength : season with freq > MUF

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
WINTER	1187	8.65	22.53	.015	2.221	.194	1.076	114.700	-44.33	2.465	83	3	1
SPRING	735	9.46	24.89	.006	1.981	-.095	.816	104.100	-44.95	1.982	75	3	1
SUMMER	775	7.35	20.66	.054	2.059	.180	1.124	107.800	-40.70	2.131	73	3	3
FALL	797	11.12	23.86	.011	2.179	-.159	1.030	117.400	-51.34	1.993	76	3	1

TABLE 14. Johnson Curve Parameters for HFBC84 Field Strength : season with freq <= MUF

condition	size	mean	sdev	beta1	beta2	gamma	delta	lambda	psi	ratio	k	t	m
WINTER	1704	-1.25	12.93	.120	4.178	-.523	2.310	26.380	-7.87	2.872	55	2	1
SPRING	2287	1.88	11.42	.010	3.706	-.209	2.682	28.440	-.50	2.311	53	2	1
SUMMER	2327	3.77	10.87	.105	3.142	8.258	5.750	401.200	-73.88	2.685	52	3	1
FALL	2465	2.45	12.11	.034	3.325	-.953	3.975	45.270	-8.85	2.639	55	2	1

Start here for all pages after the first

On first page type title of paper here

Author
Affiliation
City, State

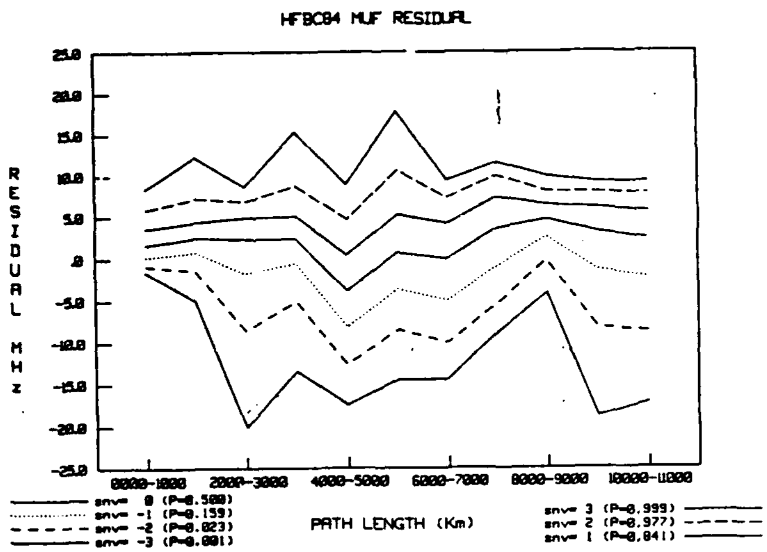


Figure 2. Predicted residual for the MUF model for the given standard normal deviates and their corresponding probability levels with path range.

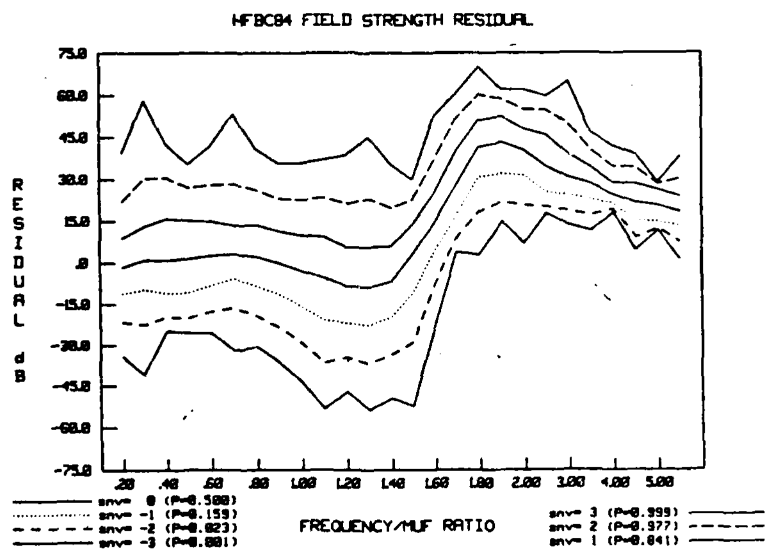


Figure 3. Predicted residual for the field strength model for the given standard normal deviates and their corresponding probability levels with f/MUF.

Leave Blank

Author's Name

Start here for all pages after the first

On first page type title of paper here

Author Affiliation City, State

DISCUSSION

A model of the probability distribution representing the residuals (errors) between the observed parameters and the corresponding predicted MUF and field strength values from HFBC84 was presented. The residuals for this program were fit to the Johnson system of frequency curves.

In the cases of both the S_U and S_B curve, the alternate method for determining the desired parameters when all four are unknown involves solving four equations representing the normal standard deviate for four percentile points. These systems of equations are non-linear, and an algorithm needs to be developed to solve these equations for the required parameters. Further, the appropriate percentile points to be used in the analysis needs to be determined.

Finally, the chi-square test of fit employed is affected by the grouping of data at particular values. At precisely the values at which a propagation model is most accurate is where the grouping occurs. This is due in part to the lack of enough precision or enough significant digits in the observed data used to evaluate the prediction model. The chi-square test of fit is used because it can be used when data from a sample is used to generate a probability distribution function. Other well known tests of fit methods can not be employed in this case.

REFERENCES

Beasley, J.D. and S.G. Springer (1977), "Algorithm AS 111. The percentage points of the normal distribution," Appl. Statist., 26, 118-121.

Dodgson, J.H. and I.D. Hill (1983), "A remark on algorithm 100: Normal-Johnson and Johnson-Normal transformations," Appl. Statist., 32, 345.

Goldberg, H. and H. Levine (1945), "Approximate formulas for the percentage points and normalization of t and χ^2," Ann. Math. Statist., 17, 216-225.

Griffiths, P. and I.D. Hill (1985), Applied Statistics Algorithms, Ellis Horwood Limited, London, 1st Ed.

Hahn, G.J. and S.S. Shapiro (1967), Statistical Models in Engineering, John Wiley and Sons, New York, 302-308.

Hill, I.D. (1973), "The normal integral," Appl. Statist., 22, 424-427.

Hill, I.D. (1976), "Algorithm AS 100. Normal-Johnson and Johnson-Normal transformations," Appl. Statist. 25, 190-192.

Hill, I.D., R. Hill and R.L. Holder (1976), "Fitting Johnson curves by moments," Appl. Statist., 25, 180-189.

Hill, I.D. and R.E. Wheeler (1981), "A remark on algorithms AS 99: fitting Johnson curves by moments and AS 100: Normal-Johnson and Johnson-Normal transformations," Appl. Statist., 30, 106.

Johnson, N.L. (1949), "Systems of frequency curves generated by methods of translation," Biometrika," 36, 149-176.

Roy, T.N. and D.B. Sailors (1987), HF maximum usable frequency (MUF) model uncertainty assessment, Naval Ocean Systems Center TR to be published, San Diego, CA.

Sprague, R.A. (1987), Accuracy of HF skywave field strength models: PROPHET and HFBC84, Naval Ocean Systems Center TR 1147, San Diego, CA.

Sprague, R.A. and D.B. Sailors (1987), Accuracy of MUF and field strength predictions by HFBC84, paper given at Fifth International Ionospheric Effects Symposium, Springfield, VA, May 5-7.

Williams, Jr., C.A. (1950), "On the choice of the number and width of classes for the chi-square test of goodness of fit," Jour. Am. Sta. Assn., 45, 77-86.

Zar, J.H. (1978), "Approximations for the percentage points of the chi-squared distribution," Appl. Statist., 27, 280-290.

Leave Blank

EMPIRICAL MODELS REPRESENTING THE ERROR IN THE PREDICTGED MUF AND FIELD STRENGTH FROM HFBC 84

David B. Sailors
Ocean and Atmospheric Sciences Division
Naval Ocean Systems Center
San Diego, CA 92152-5000

Models of the probability distribution representing the residuals between the observed MOF and field strength and the corresponding predicted values from HFBC 84, an ionospheric prediction program developed for the Broadcast WARC, is presented. A data base of 13,054 hours of oblique sounder MOFs measured on 70 paths was used to obtain the residuals in the predicted MUFs. A modified version of CCIR Data Base C was used to obtain the residuals in predicted field strength; only the 81 paths for short path propagation were retained. The residuals for these models were fit to the Johnson system of frequency curves (N. L. Johnson, Biometrika 36, 149-176, 1949). This distribution was chosen because the transformation of the Johnson variables to the normal system allows use of normal probability algorithms in its application and because it can be used to represent all univariate distributions. The Johnson system of curves consists of three distributions: (1) data that is unbounded, called the S_U distribution; (2) data that is bound on one end, called the S_L distribution; and (3) data that is bound on both ends, called the S_B distribution.

The residuals for these models were fit using an algorithm due to Hill et al which used the methods of moments (Hill et al, Appl. Statist., 25, 180-189, 1976). Since, when moments are large, the method of moments is not always efficient, the alternate methods of maximum likelihood and quantities estimation were used to obtain additional sets of Johnson parameters. Then a chi-square test of fit for a 5% level of significance was used to determine the best fit to the data.

Overall for the MUF residuals, the data can be represented by a S_L distribution. Parameters were also found for: (1) data recorder type, (2) path length and orientation, (3) season, (4) month, (5) latitude, (6) sunspot number, (7) diurnal trends, (8) geographic region, and (9) particular paths.

Overall for the field strength residuals, a S_U distribution fit the data for frequencies below or equal the predicted MUF, and a S_B distribution fit the data for frequencies greater than the MUF. The parameters were also found as a function of (1) frequency/MUF ratio, (2) particular paths in data base, (3) path length, (4) season, (5) sunspot number, (6) mid-path local time, and (7) latitude at path mid-point.

END
DATE
FILMED
5-88
DTIC